Spurious forecasts?

Martínez-Rivera, B.; Ventosa-Santaulària, Daniel; Vera-Valdés, J. Eduardo

Published in:
Journal of Forecasting

DOI (link to publication from Publisher):
10.1002/for.1219

Publication date:
2012

Document Version
Accepted author manuscript, peer reviewed version

Spurious forecasts?

Berenice Martínez-Rivera* Daniel Ventosa-Santaulària† J. Eduardo Vera-Valdés‡

This version: May 1, 2010

Abstract

P.C.B. Phillips (1998) showed that deterministic trends are a valid representation of an otherwise stochastic trending mechanism; he remained however sceptic about the predictive power of such representations. In this paper we prove that forecasts built upon spurious regression may perform (asymptotically) as well as those issued from a correctly specified regression. We derive the order in probability of several in-sample and out-of-sample predictability criteria ($\mathcal{F}$ test, Root Mean Square Error, Theil's inequality tests and R^2) using forecasts based upon a least squares-estimated regression between independent variables generated by a variety of empirically-relevant data generating processes. It is shown that, when the variables are mean stationary or trend stationary, the order in probability of these criteria is the same whether the regression is spurious or not. Simulation experiments confirm our asymptotic results.

1 Introduction

Forecasting can be considered as an activity inherent to almost any decision-making process. In economics, forecasts are commonly built upon econometric models frequently estimated by Least Squares (LS). That said, macroeconomics forecasts involve a special difficulty; many macro variables collected as time series observations appear to be nonstationary; statistical inference in models using such variables could then be misleading,

*Universidad de Guanajuato. e-mail: bbmrivera@gmail.com

†Corresponding author. Universidad de Guanajuato; Address: Departamento de Economía y Finanzas, Campus Guanajuato, Sede Marfil; Col. El Establo, 36250; e-mail: daniel@ventosa-santaularia.com

‡Centro de Investigación y Docencia Económicas. e-mail: eduardo@cimat.mx

because of the well-known spurious regression phenomenon (Yule 1926, Granger and Newbold 1974). Spurious regression has been well documented in econometrics since Phillips's (1986) provided the asymptotic arguments that explain it. It occurs whether the variables are—independently—generated as driftless unit roots (Phillips 1986), unit roots with drift (Entorf 1997), integrated of higher order processes (Marmol 1995, Marmol 1996), long-memory processes (Cappuccio and Lubian 1997, Marmol 1998) or (broken)-trend stationary processes (Hassler 2000, Noriega and Ventosa-Santaulària 2006, Noriega and Ventosa-Santaulària 2007).[1]

Notwithstanding the foregoing evidence, Phillips (1998) showed that deterministic trends are a valid representation of an otherwise stochastic trending mechanism; he remains however sceptic about the predictive power of such representations. In this paper we prove, by means of an extension of Noriega and Ventosa-Santaulària's (2007) asymptotic results, that forecasts built upon spurious regression may share (asymptotically) the properties of those issued from a correctly specified regression. This we do by proving that many of the most popular in-sample and out-of-sample predictability criteria (PrCr, hereinafter) behave asymptotically in the same manner whether the inference is spurious or not.

2 Spurious forecasts?

In words of Phillips (1998,p.1300): *"In a prototypical spurious regression, the estimated parameters are statistically significant when there is no true statistical relationship between the dependent variable and the regressors."* This phenomenon occurs whether the trending mechanism of the variables is deterministic or stochastic.

We characterize the asymptotic behavior of five classical PrCr (R^2, $\mathcal{F}$ statistic, RMSE and two versions of Theil's inequality test,[2] U_1 and U_2). Our results somehow support Phillips's (1998) conjecture: we can not correctly forecast using spurious regressions when dealing with integrated processes; nonetheless, we can do it when the series have a deterministic component .

Our study focuses in estimating a simple linear specification:

$$y_t = \alpha + \beta x_t + \epsilon_t \tag{1}$$

[1]See Ventosa-Santaulària (2009) for a more complete review of the literature concerning spurious regression.

[2]Theil (1961) proposed an error measure usually referred as"U" statistic and later ammended it (Clements and Hendry 1998, p. 63). To avoid confusion we will refer to each one with a different subscript.

Table (1) summarizes the data generating processes (DGPs) considered in this study for both the dependent and the explanatory variable:

Case	Acronym	Model
1	MS	$z_t = \mu_z + u_{zt}$
2	$I(1)$	$z_t = z_{t-1} + u_{zt}$
3	$I(1) + dr$	$z_t = Z_0 + z_{t-1} + u_{zt}$
4	TS	$z_t = \mu_z + \beta_z t + u_{zt}$
5	BTS	$z_t = \mu_z + \beta_z t + \gamma_z DT_{zt} + u_{zt}$
6	Rel^*	$y_t = \mu_y + \beta_y x_t + u_{yt}$

Table 1: The DGP's for $z_t = y_t, x_t$. *: Rel accounts for "Related variables"

In table (1) we assume that u_{xt} and u_{yt} have the following properties: $E(u_z) = 0$ and $E(u_z u_z') = \sigma^2 I$ where $u_z' = (u_{z1}, u_{z2}, \ldots u_{zT})$; $z = x, y$. DT_{zt} is a dummy variable allowing changes in the trend's slope, that is, $DT_{zt} = (t - T_{b_z}) 1 (t > T_{b_z})$, where $1(\bullet)$ is the indicator function, and T_{b_z} is the unknown date of the break in z. We denote the break fraction as $\lambda_z = (T_{b_z}/T) \in (0, 1)$ where T is the sample size.
The five in-sample/out-of-sample PrCr are enumerated below. In order to encompass both types (in and out-of sample) of PrCr, we define $\kappa = \begin{cases} t+1 \; for \; out-of \; sample \; forecasts \\ 1 \; for \; in-sample \; forecasts \end{cases}$.

1. The Root Mean Squared Error (RMSE): the RMSE is the square root of the mean of squared deviations where y_t represents the observed variable while $\hat{y}_t$ represents its corresponding prediction (see equation 2). In particular, for the in-sample version, the RMSE is just the standard deviation of the estimated residuals.

$$RMSE = \sqrt{\frac{1}{h} \sum_{i=\kappa}^{h} (y_i - \hat{y}_i)^2} \tag{2}$$

2. U Statistic U_1: Originally proposed by Theil(1958):

$$U_1 = \frac{\sqrt{\frac{1}{h} \sum_{i=\kappa}^{h} (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{h} \sum_{i=\kappa}^{h} \hat{y}_i^2} + \sqrt{\frac{1}{h} \sum_{i=\kappa}^{h} y_i^2}} \tag{3}$$

3. U Statistic U_{66}: Theil modified his own proposal; the second version of Theil's criterion, we label U_{66}, appears below:

$$U_{66} \quad = \quad \frac{\sqrt{\frac{1}{h}\sum_{i=\kappa}^{h}\left(y_i - \hat{y}_i\right)^2}}{\sqrt{\frac{1}{h}\sum_{i=\kappa}^{h} y_i^2}} \tag{4}$$

Note that Theil proposed two different formulae at different times both labelled "Coefficient of Inequality". Nowadays the U_1 is conceived as a measure of forecast accuracy whilst U_{66} is seen as a measure of forecast quality. Both measures (U_1 and U_{66}) are bounded between 0 (the case of perfect forecasting) and 1 (the supposedly worst case). Further details can be found in Bliemel (1973).

4. The $\mathcal{F}$-tests and the R^2 are frequently used in the comparison of statistical models as criteria to identify which model fits better. They can be considered as exclusively in-sample PrCr.

$$\mathcal{F} \quad = \quad \frac{\sum_{t=1}^{T}\left(\widehat{y}_t - \bar{y}\right)^2/(K-1)}{\sum_{t=1}^{T}\left(y_t - \bar{y}\right)^2/(T-K)} \tag{5}$$

$$R^2 \quad = \quad 1 - \frac{\sum_{t=1}^{T}\left(y_t - \widehat{y}_t\right)^2}{\sum_{t=1}^{T}\left(y_t - \overline{y}\right)^2} \tag{6}$$

PrCr can be therefore conducted based on (i) the in-sample fit of a model; (ii) the out-of-sample fit obtained from a sequence of recursive or rolling regressions. Inoue and Kilian (2004) overturn the conventional wisdom that out-of-sample PrCr results are more reliable than in-sample PrCr by showing that, in many cases, in-sample PrCr have higher power than out-of-sample PrCr even in the presence of parameter instability. Any out-of-sample analysis based on sample-splitting involves a loss of information and hence lower power in small samples. As a result, an out-of-sample test may fail to detect predictability that exists in the data, whereas the in-sample PrCr will correctly detect it. The previous argument lead us to also focus on in-sample PrCr.

3 Asymptotic Results

In this section we study the properties of the PrCr under a set of trending mechanisms in the data. We focus on the asymptotic behavior of the PrCr of both, in and out-of-sample forecasts.

3.1 In-Sample Forecasts

We start our study with a lemma that will prove useful along the rest of this work; it states the order in convergence of the residuals sum of squares (RSS) and the total sum of squares (TSS)[3] of equation 1's LS estimates for different trending mechanisms in the series.

Lemma 1 *Let x_t be generated by DGP $i = 1, \ldots, 5$ of table 1, let y_t be generated by DGP $j = i, 6$ and use them to estimate by LS specification (1). The order in probability of the RSS and the TSS are as follows:*

DGPs	RSS					TSS				
y_t \ x_t	MS	$I(1)$	$I(1)+dr$	TS	BTS	MS	$I(1)$	$I(1)+dr$	TS	BTS
Rel	$O_p(T)$	$O_p(T)$	$O_p(T)$	$O_p(T)$	$O_p(T)$	$O_p(T)$	$O_p(T^2)$	$O_p(T^2)$	$O_p(T^3)$	$O_p(T^3)$
$SDGP^*$	$O_p(T)$	$O_p(T^2)$	$O_p(T^2)$	$O_p(T)$	$O_p(T^3)$	$O_p(T)$	$O_p(T^2)$	$O_p(T^2)$	$O_p(T^3)$	$O_p(T^3)$

Table 2: RSS and TSS orders in convergence. *.- $SDGP$ stands for same independent DGP as x_t.

Proof: see the appendix.
On the one side, as the above table shows, the asymptotic results for the regressions using variables generated by DGPs 1 and 4 (that is, MS and TS) are qualitatively the same whether the dependent variable is related to x_t or has been independently generated by the same DGP of x_t. On the other side, the orders in probability are considerably different between regressions estimated with independent variables and those estimated with related ones when the trending mechanism is stochastic. The PrCr are particularly affected by these differences, as states proposition 1.

Proposition 1 *Let x_t and y_t be generated by DGP $i = 1, \ldots, 5$ and by DGP $j = i, 6$ of table 1 respectively, and use them to estimate by LS specification (1). The order in probability of the in-sample PrCr: $RMSE$, U_1, U_{66}, $\mathcal{F}$ and R^2, defined in (2), (3), (4), (5) and (6), respectively, are the following:*

[3] $RSS = \sum_{i=1}^{T} \left(y_{t+i} - \widehat{y_{t+i}}\right)^2$; $TSS = \sum_{t=1}^{T} \left(y_t - \overline{y}\right)^2$.

x_t	y_t	$RMSE$	U_1, U_{66}	$\mathcal{F}$	$1-R^2$
MS	Rel	$O_p(1)$	$O_p(1)$	$O_p(T)$	$O_p(1)$
	SDGP	$O_p(1)$	$O_p(1)$	$O_p(1)$	$O_p(1)$
$I(1)$	Rel	$O_p(1)$	$O_p\left(T^{-1/2}\right)$	$O_p(T^2)$	$O_p(T^{-1})$
	SDGP	$O_p\left(T^{1/2}\right)$	$O_p(1)$	$O_p\left(T^{3/2}\right)$	$O_p(1)$
$I(1)+dr$	Rel	$O_p(1)$	$O_p\left(T^{-1/2}\right)$	$O_p(T^2)$	$O_p(T^{-1})$
	SDGP	$O_p\left(T^{1/2}\right)$	$O_p(1)$	$O_p\left(T^{3/2}\right)$	$O_p(1)$
TS	Rel	$O_p(1)$	$O_p(T^{-1})$	$O_p(T^3)$	$O_p(T^{-2})$
	SDGP	$O_p(1)$	$O_p(T^{-1})$	$O_p(T^3)$	$O_p(T^{-2})$
BTS	Rel	$O_p(1)$	$O_p(T^{-1})$	$O_p(T^3)$	$O_p(T^{-2})$
	SDGP	$O_p(T)$	$O_p(1)$	$O_p(T)$	$O_p(1)$

Proof: see the appendix.

Proposition 1 shows that, when x_t and y_t are either mean stationary or trend stationary, the order in probability of the three in-sample PrCr ($RMSE$, U_1 and U_{66}) are the same whether the variables are independent or not. Furthermore, we should take a more careful approach when we are analysing integrated processes; Proposition 1 proves that the order in probability of the statistics are different between related variables and between independent ones; forecasts based on the former variables are more accurate than those based on the latter variables. This in line with the conjecture made by Phillips (1998). In words of Phillips (1998,p.1300): *"It is important to recognize that such representations* **[That is, a spurious regression]** *(. . .) do not take the place of temporal predictive models."*

Furthermore, results in lemma 1 and proposition 1 are consistent with the idea that, when dealing with processes governed by a deterministic trend (as long as there are no structural breaks), the PrCr have the same order in probability whether the variables are related or not. However, in the presence of structural breaks, the spurious regression losses its prediction capability. The issue of structural breaks is not a minor one. To quote Pesaran, Pettenuzzo, and Timmermann (2006): "*Structural changes or "breaks" appear to affect models for the evolution in key economic and financial time series such as output growth, inflation, exchange rates, interest rates and stock returns. This could reflect legislative, institutional or technological changes, shifts in economic policy, or could even be due to large macroeconomic shocks such as the doubling or quadrupling of oil prices experienced over the past decades. . .* ". We further study the presence of structural breaks in the DGP. In proposition 2 we present the order in probability of the PrCr when the nonsense regression correctly specifies the structural break; this specification is represented in eq.

7.[4]

$$y_t = \alpha + \beta x_t + \gamma DT_{yt} + \epsilon_t \tag{7}$$

Proposition 2 *Let x_t and y_t be generated by DGP 4 and 5 of table 1, respectively; use both to estimate by LS equation (7). Hence, the orders in probability of the PrCr: RSS, TSS, $RMSE$, U_1 and U_{66}, $\mathcal{F}$ and R^2, are the following:*

RSS	TSS	$RMSE$	U_1, U_{66}	$\mathcal{F}$	$1-R^2$
$O_p(T)$	$O_p(T^3)$	$O_p(1)$	$O_p(T^{-1})$	$O_p(T^3)$	$O_p(T^{-2})$

Proof: see the appendix.
A comparison between the results of proposition 2 with those of proposition 1 and lemma 1 reveals striking similarities. The asymptotic behaviour of the PrCr is the same, whether the variables are related or not, as long as the structural break has been correctly modelled. Our results might help to better understand the properties of a *spurious regression*. The commonality of a deterministic trending mechanism between two otherwise independent variables allows for a suitable linear combination of both variables capable of providing an useful tool in the forecasting of the long term evolution of one of the variables. That said, the later statement certainly does not entail any causal link between the variables.
In other words, although the variables could clearly be economically linked, we could also explain the long horizon predictability capacities of the regressor in terms of a common type of trend. Nevertheless, further investigation should be done concerning which data generating processes are more appropriate in the variables's forecasting. This would allow us to determined under which circumstances the attractive properties of the *spurious forecasts* hold. Yet, when a deterministic trending mechanism is preferred by statistical evidence, then one should consider the possibility that whether the specification is "*spurious*" or not, forecasts do have similar properties. Moreover, if there is a structural break in the dependent variable but the *spurious* specification includes it, the forecasting capacity of such spurious regression are maintained.

3.2 Out-of-Sample Forecasts

We turn now our attention to out-of-sample forecasts. The main difference between an in-sample and an out-of-sample forecast lies in the amount of information they use. An

[4]It is assumed that the structural break DT_{yt} corresponds to the one defined in case 5 (see Table 1).

in-sample forecast employs all of the available information as do the parameter estimates whilst an out-of-sample forecast only uses a fraction of the sample to estimate the parameters. To obtain the asymptotics of out-of-sample PrCr we are thus forced to split the sample size. The procedure to build q out-of-sample forecasts (where q is an integer larger than 0) in this work can be described as follows: (i) we first estimate the regression using the whole sample size (T) and obtain $\hat{\alpha}$ and $\hat{\beta}$; (ii) we then use the parameter estimates to build q forecasts of the variable y_t: $\hat{y}_{t+i} = \hat{\alpha} + \hat{\beta}x_{t+i}$ where $i = 1, 2, \ldots, q$ (note that we assume that x_{t+i} is known); (iii) thirdly, we use the true DGP of y_t to obtain the real future values, y_{t+i}, and; (iv) finally, prediction errors can then be obtained simply by calculating the forecast error: $\hat{u}_{t+i} = y_{t+i} - (\hat{\alpha} + \hat{\beta}x_{t+i})$.

We focus on the MS and the TS processes. This accounts for the fact that they both have a deterministic trending mechanism attached to them[5] so once we estimate it, we can correctly forecast the future values of the series. The main finding is that on both regressions (the one that uses independent variables as well as the one that uses related variables), we obtain the same order in probability for the PrCr. This is proved in Proposition 3.

Proposition 3 *Let x_t and y_t be generated by DGPs $1, 4$ and $1,\ 2$ and 6 of table 1, respectively, and use them to estimate by LS specification (1) and (7). The asymptotics of the estimated parameters $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ are as shown on table 3 below.*

DGPs	$\widehat{\alpha}$				$\widehat{\beta}$				$\widehat{\gamma}$	
x_t \ y_t	MS	TS	Rel	BTS^*	MS	TS	Rel	BTS^*	Rel	BTS^*
MS	μ_y	–	μ_y	–	0	–	β_y	–	–	–
TS	–	$-\frac{\beta_y\mu_x+\beta_x\mu_y}{\beta_x}$	μ_y	μ_y	–	$\frac{\beta_y}{\beta_x}$	β_y	β_y	γ_y	γ_y

Table 3: Asymptotic parameters
* for the TS-BTS combination the specification (7) is employed.

Proof: see the appendix.

Results in the previous proposition allows us to obtain the asymptotic behavior of the regression's residuals (see table 4):

[5] In the first case, MS, the trend is zero.

DGPs		$\widehat{u}_q$		
x_t \ y_t	MS	TS	Rel	BTS^*
MS	u_{yq}	–	u_{yq}	–
TS	–	$-\frac{\beta_y}{\beta_x}u_{xq} + u_{yq}$	u_{yq}	$-\frac{\beta_y}{\beta_x}u_{xq} + u_{yq}$

Table 4: Asymptotic parameters
*The TS-BTS combination uses specification (7)

The results from proposition 3 show that the prediction errors yielded by a regression that uses independent variables are much in line with those obtained when the variables are related. Moreover, it can also be seen that all the PrCr analyzed in this work have the same asymptotic behavior whether they are in-sample or out-of-sample. Yet, a small difference should be noted when the variables are generated as DGP 4: the forecasts have a larger variance if the variables are independent (the difference accounts for the term u_{xq}). Nonetheless the forecast-error mean remains zero.
Finally, when working with DGP 5 (a BTS process), the PrCr have the same asymptotic behavior as in the previous section, as long as the structural break is correctly modelled.

4 Finite-sample evidence

We compute the value of the RMSE, U_1 and U_{66} PrCr, using simulated data, for various sample sizes and DGPs combinations. Our results appear in tables 5 and 6.

Sample	Combinations of cases ($DGPs$) in table 1								
T	$MS-MS$	$Rel-MS$	$TS-TS$	$Rel-TS$	$I(1)-I(1)$	$Rel-I(1)$	$BTS-BTS$	$Rel-BTS$	$I(1)-TS$
50	0.994	0.996	1.403	0.994	1.892	0.995	3.478	0.994	3.158
100	0.997	0.995	1.409	0.997	2.630	0.996	6.300	0.997	5.423
250	0.999	0.998	1.413	0.999	4.175	0.999	15.14	0.999	12.42
500	0.999	0.999	1.413	0.999	5.844	0.999	30.02	0.999	24.15
1000	0.999	0.999	1.413	0.999	8.247	0.999	59.84	0.999	47.79

Table 5: Average of RMSE statistic for $10,000$ simulations

From table 5, it is straightforward to see, on the one hand, that, when the variable is either TS or MS, the $RMSE$ behaves in the same manner whether the variables are related or not. On the other hand the values of columns 6,8,10 increase as the sample size grows, so the $RMSE$ is capable of detecting a spurious regression when the variables are a unit root with drift or a broken trend stationary. In table 6 we compute the U_1 statistic using

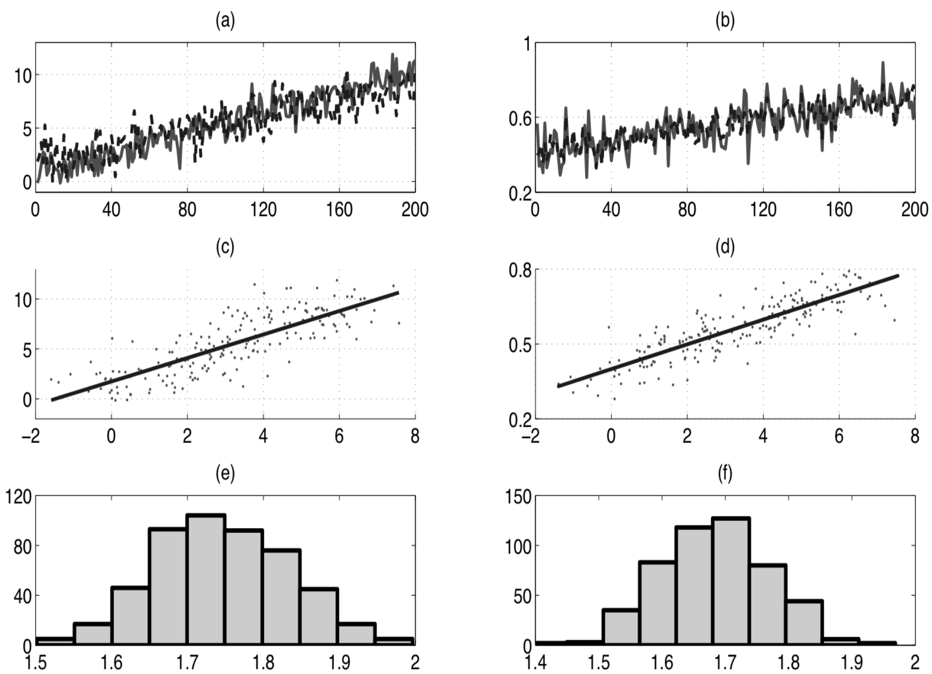


Figure 1: In-sample forecasts: In figures (a) and (b) the pointed line represents the forecast accuracy achieved by a spurious (correctly specified) regression where x_t and y_t are both generated by a TS DGP (x_t is generated by a TS DGP while y_t is generated by the true DGP) while the continuous line accounts for the actual value. Time is represented in the X axis whilst the Y axis indicates the value of y_t. In figures (c) and (d) the line is the in-sample forecasts for each value of the variable x_t; the scatter plot is also shown. The X and Y axes represent, the value of x_t and y_t, respectively. In figures (e) and (f) histograms for the $RMSE$ in-sample forecast statistic are presented.

simulated data (results for the U_{66} are available upon request). It can also be seen that the value of this PrCr is close to zero when the variable is trend stationary. This indicates an acceptable forecasting whether the regression is spurious or not, even for an small value of T. In the case of a mean stationary variable the U_1 statistic value is not quite close to zero, but either way the statistic is not capable of detecting an incorrect specification. In figure 1 it can be seen that, when the variables are generated by a TS process, the accuracy achieved by LS forecasts is similar, whether the variables used in the LS regression are

Sample	Combinations of cases $(DGPs)$ in table 1								
T	$MS-MS$	$Rel-MS$	$TS-TS$	$Rel-TS$	$I(1)-I(1)$	$Rel-I(1)$	$BTS-BTS$	$Rel-BTS$	$I(1)-TS$
50	0.207	0.207	0.015	0.005	0.032	0.002	0.030	0.004	0.025
100	0.207	0.207	0.008	0.002	0.022	0.001	0.028	0.002	0.022
250	0.207	0.207	0.003	0.001	0.014	0.000	0.027	0.001	0.020
500	0.208	0.207	0.001	0.000	0.010	0.000	0.027	0.000	0.019
1000	0.208	0.208	0.000	0.000	0.007	0.000	0.027	0.000	0.019

Table 6: Average of U_1 statistic for $10,000$ simulations

independent or not. Figure 1 illustrates the asymptotic results; note that the histograms for the in-sample $RMSE$ are also alike, whether the regression is spurious or not.
Finally, in figure 2 we plot the residual sum of squares of both, the spurious regression and the correctly specified one for the TS-DGP case. We allow for different values of the variance of the processes as well as a small covariance between their innovations. As the figure shows (in accordance to proposition 3 results), as the quotient of the parameters increases $(\frac{\beta_y}{\beta_x})$ so does our errors on the forecasts. Nonetheless, the PrCr on both regressions behave asymptotically the same.

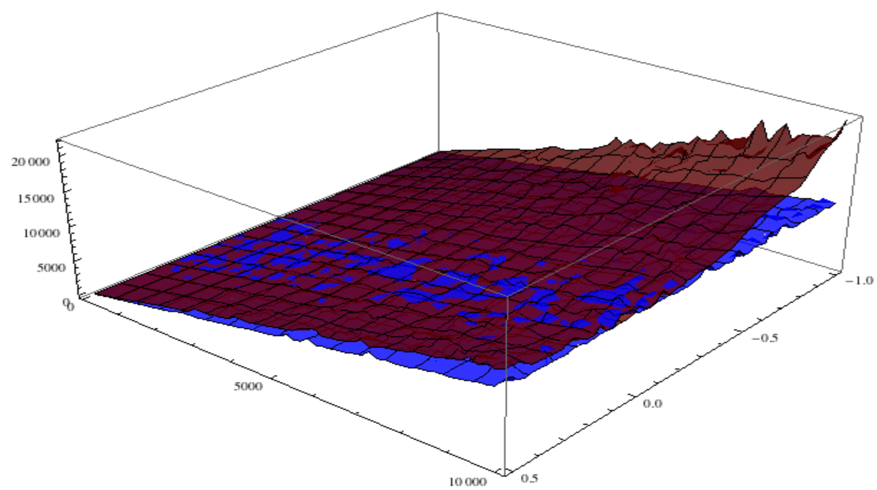
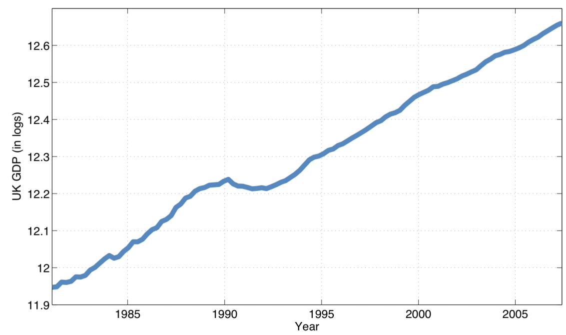


Figure 2: RSS statistic on a spurious (darker blanket) and a correctly specified regression (lighter one) as the size of the sample increases.

5 Forecasting the GDP of UK

We illustrate our previous results using real data of the United Kingdom GDP. The dataset ranges from the first quarter of 1981 to the fourth quarter of 2007.[6] We use the dataset to estimate, by LS, a regression where the log of the UK GDP is the dependent variable whilst the regressor is a simulated variable. This we do to ensure the independence between the variables. According to our previous results, the spurious forecasts do have desirable properties when the regressor, as well as the regressand have been generated as TS or BTS process. In the case of the simulated variable, x_t has been simulated according to the following TS DGP:



$$x_t \quad = \quad 12 + 0.01 \cdot t + u_t \tag{8}$$

Where $u_t \sim \mathcal{N}\left(0, (1/400)^2\right)$ and $t = 1, 2, \ldots, 108$.
As for the dependent variable, we typify the trending mechanism of the series using Perron's (1997) (P97, hereinafter) test. We employed P97's second auxiliary regression, which corresponds to the *Changing Growth* model (Perron 1989). The resulting statistic yielded a value of -6.2646 and thus provided enough evidence to reject the null hypothesis (using the critical value at a 1% level) of a unit root process in favor of a trend stationary process. The P97 test also detected a break in the second quarter of 1990. This informations allows us to fully specify the "spurious regression" as in eq. 7. Note that, in order to compute out-of-sample forecasts, we can only use a subsample of the entire dataset. The LS regression has been therefore estimated using only observations from the first quarter of 1981 to the third quarter of 2001:

$$\begin{array}{lclllll} log(GDP_{UK})_t & = & \alpha \ + & \beta x_t \ + & \gamma_1 DU_t \ + & \gamma_2 DT_t & + \epsilon_t \\ & & 1.41 & 0.87 & -0.07 & -0.002 & \\ & & (5.16) & (39.04) & (-11.36) & (-6.54) & \end{array} \tag{9}$$

[6]Source: OCDE statistics, available at the Mexican Statistics Office (INEGI) website: `www.inegi.gob.mx` The GDP is at constant prices of 2003 and has been seasonally adjusted.

Where DU_t accounts for the change in level and DT_t, for the change in trend. Note that the R^2 is quite high (0.9918) and the classical $\mathcal{F}$ statistic (26.4489) rejects the null hypothesis of $\beta = \gamma_1 = \gamma_2 = 0$ which is in line with what is known about spurious regression between BTS processes (see Noriega and Ventosa-Santaulària, 2007). We next forecast the UK GDP series using the estimates of the spurious regression 9. As previously pointed, we only employed the first 84 observations and built upon them out-of-sample forecasts with the resulting horizon $\kappa = 24$. Figure 5 shows the real and the predicted values of the series. The forecasts based upon a "spurious regression" are quite accurate, which is in line with our previous results concerning deterministically-trended data. Note that Theil's inequality PrCrs support this statement: the computed U_1 and U_2 PrCr yielded the following statistics values: 0.0002 and 0.0003, respectively.

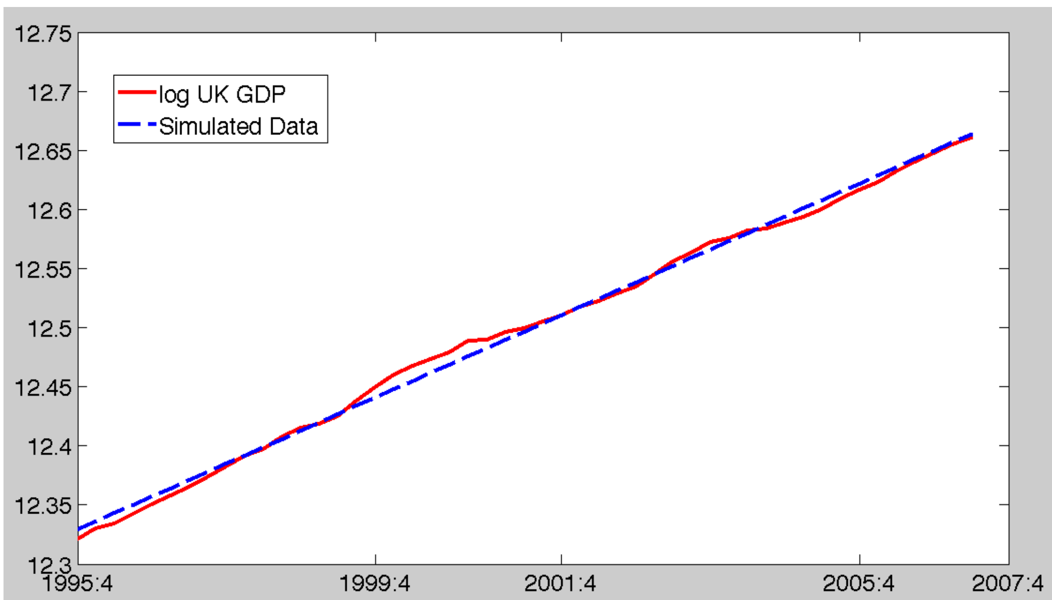


Figure 3: Real and simulated data for the UK GDP

The previous statistical exercise should prove useful in the prediction of the UK real GDP, as long, of course, as there are not any other undetected structural break in the near future (such as the one that had may occurred in 2008-2009).

6 Concluding remarks

This paper proves that, when two independent variables share the same type of deterministic trending mechanism, several widely employed in-sample and out-of-sample predictability criteria behave asymptotically in the same manner. In particular, we showed that, when the variables are stationary, either around a constant or a trend, the behaviour of in-sample and out-of-sample PrCr is similar whether the relation between the variables exists or the variables are independent from each other. It is also shown that even a spurious regression with a correctly modelled structural break provides adequate forecasts. Our results also support Phillips's conjecture since they show that forecasts based upon independent integrated variables compare poorly to the ones obtained from an estimated cointegrated regression.

A Proof of Lemma 1 and Propositions 1, 2 and 3

We present a guide as to how to obtain the order in probability of the five predictability statistics appearing in Proposition 2 in the estimation of regression (7) by LS where the variables y and x are generated by DGP 5 and DGP 4 respectively (all other combinations follow the same steps. Proof of such was provided with the aid of Mathematica 4.1 software). We use the classical LS formulas:

$$\widehat{\beta}_{LS} = (X'X)^{-1} X'Y$$

where

$$X'X = \begin{bmatrix} T & \sum x_t & \sum DT_{yt} \\ \sum x_t & \sum x_t^2 & \sum x_t DT_{yt} \\ \sum DT_{yt} & \sum x_t DT_{yt} & \sum DT_{yt}^2 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} \sum y_t \\ \sum x_t y_t \\ \sum DT_{yt} y_t \end{bmatrix}$$

Unless otherwise indicated all sums run from $t = 1$ to T

$$\sum x_t = \mu_x T + \beta_x \sum t + \underbrace{\sum u_{xt}}_{O_p\left(T^{\frac{1}{2}}\right)}$$

$$\sum y_t = \mu_y T + \beta_y \sum t + \gamma_y \sum DT_{yt} + \underbrace{\sum u_{yt}}_{O_p\left(T^{\frac{1}{2}}\right)}$$

$$\begin{aligned}\sum x_t^2 = {} & \mu_x^2 T + \beta_x^2 \sum t^2 + \underbrace{\sum u_{xt}^2}_{O_p(T)} + 2\mu_x\beta_x \sum t + 2\mu_x \underbrace{\sum u_{xt}}_{O_p\left(T^{\frac{1}{2}}\right)} \\ & + 2\beta_x \underbrace{\sum t u_{xt}}_{O_p\left(T^{\frac{3}{2}}\right)}\end{aligned}$$

$$\begin{aligned}\sum y_t^2 = {} & \mu_y^2 T + \beta_y^2 \sum t^2 + \gamma_y^2 \sum DT_{yt}^2 + \underbrace{\sum u_{yt}^2}_{O_p(T)} + 2\mu_y\beta_y \sum t \\ & + 2\mu_y\gamma_y \sum DT_{yt} + 2\mu_y \underbrace{\sum u_{yt}}_{O_p\left(T^{\frac{1}{2}}\right)} + 2\beta_y\gamma_y \sum tDT_{yt} \\ & + 2\beta_y \underbrace{\sum t u_{yt}}_{O_p\left(T^{\frac{3}{2}}\right)} + 2\gamma_y \underbrace{\sum DT_{yt} u_{yt}}_{O_p\left(T^{\frac{3}{2}}\right)}\end{aligned}$$

$$\begin{aligned}\sum y_t x_t = {} & \mu_x\mu_y T + \mu_x\beta_y \sum t + \mu_x\gamma_y \sum DT_{yt} + \mu_x \underbrace{\sum u_{yt}}_{O_p\left(T^{\frac{1}{2}}\right)} \\ & + \beta_x\mu_y \sum t + \beta_x\beta_y \sum t^2 + \beta_x\gamma_y \sum tDT_{yt} + \beta_x \underbrace{\sum t u_{yt}}_{O_p\left(T^{\frac{3}{2}}\right)} \\ & + \mu_y \underbrace{\sum u_{xt}}_{O_p\left(T^{\frac{1}{2}}\right)} + \beta_y \underbrace{\sum t u_{xt}}_{O_p\left(T^{\frac{3}{2}}\right)} + \gamma_y \underbrace{\sum DT_{yt} u_{xt}}_{O_p\left(T^{\frac{3}{2}}\right)} + \underbrace{\sum u_{yt} u_{xt}}_{O_p\left(T^{\frac{1}{2}}\right)}\end{aligned}$$

$$\sum x_t DT_{yt} = \mu_x \sum DT_{yt} + \beta_x \sum tDT_{yt} + \underbrace{\sum DT_{yt} u_{xt}}_{O_p\left(T^{\frac{3}{2}}\right)}$$

$$\sum y_t DT_{yt} = \mu_y \sum DT_{yt} + \beta_y \sum tDT_{yt} + \gamma_y \sum DT_{yt}^2 + \underbrace{\sum DT_{yt} u_{yt}}_{O_p\left(T^{\frac{3}{2}}\right)}$$

where,

$\sum t = \frac{1}{2}\left(T^2 + T\right)$
$\sum t^2 = \frac{1}{6}\left(2T^3 + 3T^2 + T\right)$
$\sum DT_{wt} = \frac{1}{2}\left[T^2(1-\lambda_w)^2 + T(1-\lambda_w)\right]$
$\sum DT_{wt}^2 = \frac{1}{6}\left[2T^3(1-\lambda_w)^3 + 3T^2(1-\lambda_w)^2 + T(1-\lambda_w)\right]$
$\sum tDT_{wt} = \lambda_w T \sum DT_{wt} + \sum DT_{wt}^2$

The orders in convergence of the underbraced expressions can be found in Hamilton (1994) pp.505-506. and in Noriega and Ventosa-Santaularia (2007). We can fill the previously-cited matrices and then compute the LS parameters estimates and the $RMSE$, U_{58}, U_{58}, $\mathcal{F}$ and R^2 associated.
All programs in their true format and extension are available to the reader upon request.

B Simulation Parameters

Parameters used for the simulations:

References

BLIEMEL, F. (1973): "Theil's forecast accuracy coefficient: A clarification," *Journal of Marketing Research*, 10(4), 444–446.

CAPPUCCIO, N., AND D. LUBIAN (1997): "Spurious Regressions Between I(1) Processes with Long Memory Errors," *Journal of Time Series Analysis*, 18, 341–354.

	Table 5 and Table 6						Figure 1		
Parameter	1-1,5-1	3-3,5-3	2-2	5-2	4-4	5-4	a,b,c	d,e	f
σ_y	1	1	1	1	1	1	1	0.062	1.7
σ_x	1	1	1	1	1	1	1	1	1
μ_y	2.3	0.89	1	2.1	0.89	0.89	0.4	0.4	0.4
μ_x	1.5	1.7	3	3	1.7	1.7	0.2	0.2	0.2
β_y	0	1.9	1	2.5	1.9	1.9	0.05	0.05	0.05
β_x	0	1.5	1	1	1.5	1.5	0.03	0.03	0.03
γ_y	0	0	0	0	0.7	0	0	0	0
γ_x	0	0	0	0	1.15	1.3	0	0	0
λ_y	0	0	0	0	0.3	0	0	0	0
λ_x	0	0	0	0	0.7	0.6	0	0	0

CLEMENTS, M., AND D. HENDRY (1998): *Forecasting economic time series*. Cambridge University Press.

ENTORF, H. (1997): "Random Walks Wiht Drifts: Nonsense Regression and Spurious Fixed-Effect Estimation," *Journal of Econometrics*, 80, 287–296.

GRANGER, C., AND P. NEWBOLD (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 11–20.

HAMILTON, J. (1994): *Time Series Analisys*. Princeton.

HASSLER, U. (2000): "Simple Regressions with Linear Time Trends," *Journal of Time Series Analysis*, 21, 27–32.

INOUE, A., AND L. KILIAN (2004): "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?," *Econometric reviews*, (4), 371.

MARMOL, F. (1995): "Spurious Regressions Between I(d) Processes," *Journal of Time Series Analysis*, 16, 313–321.

MARMOL, F. (1996): "Nonsense Regressions between Integrated Processes of Different Orders.," *Oxford Bulletin of Economics & Statistics*, 58(3), 525–36.

MARMOL, F. (1998): "Spurious Regression Theory with Nonstationary Fractionally Integrated Processes," *Journal of Econometrics*, 84, 233–250.

NORIEGA, A., AND D. VENTOSA-SANTAULÀRIA (2006): “Spurious Regression Under Broken Trend Stationarity,” *Journal of Time Series Analysis*, 27, 671–684.

——— (2007): “Spurious Regression And Trending Variables,” *Oxford Bulletin of Economics and Statistics*, 7, 4–7.

PERRON, P. (1989): “The Great Crash, The Oil Price Shock and the Unit Root Hypothesis,” *Econometrica*, 57, 1361–1401.

——— (1997): “Further evidence on breaking trend functions in macroeconomic variables,” *Journal of Econometrics*, 80(2), 355–385.

PESARAN, M., D. PETTENUZZO, AND A. TIMMERMANN (2006): “Forecasting time series subject to multiple structural breaks,” *Review of Economic Studies*, 73(4), 1057–1084.

PHILLIPS, P. (1986): “Understanding Spurious Regressions in Econometrics,” *Journal of Econometrics*, 33, 311–340.

——— (1998): “New Tools for Understanding Spurious Regressions,” *Econometrica*, 66, 1299–1325.

THEIL, H. (1961): *Economic Forecasts and Policy*. North Holland, (1966) Applied Economic Forecasting. North Holland.

VENTOSA-SANTAULÀRIA, D. (2009): “Spurious Regression,” *Journal of Probability and Statistics*, (2009), 27 pp., doi:10.1155/2009/802975; `http://downloads.hindawi.com/journals/jps/2009/802975.pdf`.

YULE, G. (1926): “Why do we Sometimes get Nonsense-Correlations between Time-Series?–A Study in Sampling and the Nature of Time-Series,” *Journal of the Royal Statistical Society*, 89(1), 1–63.